**An Analysis on Staff and Patient Performance for Anticoagulation**

**with Intravenous Unfractionated Heparin**

Emma Stevens

Ken Jon Yeong

April 26, 2015

MT 422: Applied Statistics

John Carroll University

**TABLE OF CONTENTS**

1. INTRODUCTION………………………………………………………………...3
2. DATASET DESCRIPTION AND MODIFICATIONS…………………………..4
3. ANALYSIS OF STAFF PERFORMANCE………………………………………7
4. TIME IN THERAPEUTIC RANGE…………………….……………………....11

1. TIME TO INITIAL aPTT…………………………………………...…………...22
2. CONCLUSION…………………………………………………………………..26
3. REFERENCES…………………………………………………………………..29

**1. Introduction**

In the spring of 2015, Dr. Robert Patrick from the Louis Stokes Veteran Affairs Medical Center approached Dr. Consiglio of John Carroll University for assistance to assess the medical center’s performance of anticoagulation with intravenous unfractionated (IVUFH). This project was subsequently introduced to us through Dr. Consiglio’s applied statistics class.

Dr. Patrick was interested in two types of performance, staff performance and patient performance. Staff performance is defined by the ability of staff to accurately implement the hospital’s treatment procedure. Patient performance is defined by the quality of treatment that a patient receives. This paper is an assessment of these performances.

In Section 2, we will address the dataset that Dr. Patrick provided. We will discuss the original structure of the dataset and the modifications that were made to it for statistical analysis. In Section 3, we will analyze staff performance. We will introduce the hospital’s treatment procedure and perform an analysis to determine how well the procedure was implemented. We will also discuss dataset anomalies that we encountered while performing our analysis.

In Section 4 and 5, we will discuss patient performance through two measures of performance, time spent in therapeutic range (Section 4) and time to initial activated partial thromboplastin times (aPTT) (Section 5). We will introduce three methods to measure time spent in therapeutic range and two methods to measure time to initial aPTT. We will provide a comparison between the three methods in Section 4 and we will analyze the relationship between the two measures of performance in Section 5.

In Section 6, we will discuss the results of our analysis and future research. All analysis was done in R Studio using the base R package as well as the lubridate [[1]] and plyr [[4]] package.

**2. Dataset Description and Modifications**

The original dataset contained 52339 rows and six columns. These measurements were taken from 3899 adults with 4960 episodes. No information about a patient’s gender, age, race or health status was provided. We were only given identification for the episodes through the “Episode\_Id” column.

Each row details the collection time (“Collection\_Time” column) and the collected value (“Numeric\_Value” column) for a particular collection. On average, there were around 11 collections per episode. The other three columns (“Order\_Start”, “Order\_End”, and “Bolus”) details the data and time when treatment was ordered, treatment ended; Note, the time treatment was ordered is not the same as the time treatment was started. and whether or not an initial bolus was given at the beginning of treatment. A value of “1” indicates that an initial bolus was present and a value of NA indicates otherwise. Rows with identical ids had identical “Order\_Start”, “Order\_End” and “Bolus” values. Table 2.1 provides the rows that detailed the collection time and collected value for episode seven.

Table 2.1: Collection times and values from episode seven.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Episode**  **Id** | **Order**  **Start** | **Order**  **End** | **Bolus** | **Collection**  **Time** | **Numeric**  **Result** |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/16/10 16:59 | 180 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/17/10 2:50 | 83.7 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/17/10 11:35 | 83.9 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/17/10 21:09 | 37 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/19/10 1:45 | 58.7 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/19/10 9:00 | 61.2 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/19/10 16:30 | 78.3 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/19/10 23:44 | 68.3 |
| 7 | 10/16/10 7:30 | 10/20/10 10:11 | 1 | 10/20/10 5:30 | 97.3 |

An additional five columns (STC, OMI, TRange, EMI, Late\_by) were added to the dataset. The “Start to Collection” (STC) column provides the elapsed time (in hours) between the “Order\_Start” and “Collection\_Time” column. The “observed monitoring interval” (OMI) column provides the elapsed time between the previous collection time and the current collection time. However, for the first collection of every episode, we used the STC value for OMI. Since the “Order\_Start” time does not necessarily represent the beginning of treatment, the OMI value for these first collections might not be accurate.We will address this more in Section 3. All measurements of time are given in hour units.

The “TRange” column labels whether a measurement is above, inside or below therapeutic range. The therapeutic range is defined between and including aPTT values from 63 to 95 [[2], 6]. Hospital procedure mandates a repeat aPTT six hours after the beginning of treatment or after the previous dose. This continues until two consecutive aPTT’s are inside therapeutic range, after which the monitoring interval changes to a daily AM interval. The interval switches back to every six hours if the aPTT falls outside therapeutic range [[2], 7]. In total, 22.3, 40.0 and 37.7 percent of collections were labeled “Above”, “Inside” and “Below” respectively.

The “estimated monitoring interval” (EMI) column tells us whether the collection has a six-hour monitoring interval (6) or a daily AM monitoring interval (24). Dr. Patrick conveyed that the hospital did not have a scheduled daily AM time. Collections were taken at five o clock, seven o clock or nine o clock in the morning. Collections could have also been taken 24 hours after the previous collection. Due to time constraints, we did not have time to make the necessary changes to our dataset. Therefore we did not perform an analysis on the time difference for collections with a daily AM monitoring interval. In total, about 83 and 17 percent of collections were labeled “6” and “24” respectively.

The “Late\_by” (LB) column measures the difference between the “OMI” column and the “EMI” column. For example, in collection two, Episode seven has an EMI value of 6 since the previous aPTT was above therapeutic range. However, the collection time was 9.83 hours after the previous collection time. Thus, this collection has an LB value of 3.83 hours (9.83- 6). Note that OMI is inaccurate for collections with a daily AM monitoring interval and the first collection of each episode. We will discuss our analysis of LB in Section 3. Table 2.2 provides sample rows for the added columns.

Table 2.2: Sample row from dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Episode\_Id** | **STC** | **TRange** | **OMI** | **EMI** | **LB** |
| 7 | 9.5 | Above | 9.5 | 6 | 3.5 |

While running our analysis, we found two kinds of anomalies. The first consisted of collections with identical episode Ids and collection times but with different numeric results. The second consisted of collections with identical order start and collection times. The latter kind only affected the kind of analysis we wanted to perform in Section 5. Thus, we only excluded them in Section 5. In total there were 107 and 67 collections of the first and second kind of anomaly respectively.

After comparing each of the 107 collections (first anomalies) with the collections before it, we excluded all collections whose aPTT difference with the previous collection was less than two. We assumed these differences were negligible since they were so close to each other. For consecutive collections within two numeric points but different “TRange” values (“Above” for one collection, “Inside” or “Below” for another), we understand that removing these collections changes the percent of time in range as measured by the first and second method in Section 4. However, we see this more as an inadequacy of the methods themselves to identify and disqualify such anomalies in the dataset.

In total, 57 collections were removed. As for the other 50 collections, we eliminated them and all collections with identical episode Ids. This eliminated another 860 collections. Thus the final dataset had 51422 rows and 4910 episodes. Table 2.3 provides a sample of the anomalies encountered.

Table 2.3 Dataset Anomalies

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Row Number** | **Episode\_Id** | **Collection\_Time** | **Numeric Results** | **TRange** |
| 44611 | 14194 | 7/25/10 5:28 | 131.4 | Above |
| 44612 | 14194 | 7/25/10 5:28 | 180 | Above |
| 22296 | 7283 | 12/21/10 20:30 | 49.8 | Below |
| 22297 | 7283 | 12/21/10 20:30 | 80.2 | Inside |
| 776 | 298 | 6/15/10 7:02 | 96.9 | Above |
| 777 | 298 | 6/15/10 7:02 | 95.3 | Above |

Using these created variables, we were able to create another dataset that detailed the time spent above, inside and below therapeutic range for each episode Id. As mentioned in Section 2, three different methods were used to calculate time spent above, inside and below therapeutic range. In total there were 4910 rows (one for each episode) and 22 columns, one “Episode\_Id” column and seven columns for each method (time above, time inside, time below, total time, percent above, percent below and percent inside). Table 2.3 provides a sample of our dataset with one of the methods.

Table 2.3: Sample row from created dataset.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Episode Id** | **Rows Above** | **Rows Below** | **Rows Inside** | **Total Rows** | **FOV %Above** | **FOV %Below** | **FOV %Inside** |
| 7 | 2 | 3 | 4 | 9 | 22.2 | 33.3 | 44.4 |

**3. Analysis of Staff Performance**

In this section, we will analyze staff performance to determine how well hospital staff followed hospital procedure in the administration of IVUFH. Recall that hospital procedure mandates a repeat aPTT every six hours until two consecutive aPTTs are inside therapeutic range, after which the monitoring interval changes to a daily AM interval. The interval switches back to every six hours if the aPTT falls outside therapeutic range. In this section, we will analyze how well hospital staffs followed these procedures.

In Section 2, we introduced the LB column. We mentioned that LB was inaccurate for collections with a daily AM monitoring interval or if the collection was the first collection of an episode. Thus, in order to run an analysis on LB, we need to exclude these collections first. That left us with a remainder of 37769 collections. Table 3.1 provides a numerical summary of LB for these collections.

Table 3.1: Numerical summary of LB

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| -5.9 | 0.2 | 2 | 2.9 | 4.3 | 282.8 |

The minimum and maximum values from Table 3.1 were unexpected. Recall that LB measures the difference between OMI and EMI. Therefore if collections were collected and recorded accurately, we should expect our LB values to hover close to zero.

A time difference of -5.9 happens when the previous and current collection were only 6 minutes apart. We examined our dataset and found that 220 collections had LB values between -5 to -6. This meant that these collections were taken less than an hour after their previous collection. We are unsure why this would be the case. We examined the data but found no intelligible pattern. Perhaps some hospital personnel were unaware that a collection had been taken not too long ago, and took another collection. Perhaps another collection was ordered.

On the other hand, a maximum of 282.8 meant that 288.8 hours (about 12 days) went by before the next collection was taken. This seems like an unlikely situation since only six hours should go by. We found 135 collections that took over 24 hours. Perhaps there are missing records from these episodes. Perhaps there really is negligence on the part of hospital personnel.

The mean from Table 3.1 tells us that on average, collections were taken 2.9 hours after they were supposed to be taken. To measure the hospital’s performance, we calculate the percent of collections that were taken within one, two and three hours of the supposed time. Table 3.2 provides the percentages of collections taken within one, two and three hours of the supposed time.

Table 3.2: Percentages of collections taken within one, two and three hours of supposed time.

|  |  |
| --- | --- |
| **Hours** | **Percent** |
| ±1 | 23.0 |
| ±2 | 41.1 |
| ±3 | 57.8 |

Based on the results in Table 3.2, we see that there is room for improvement for hospital staff to follow hospital procedure. More than 40 percent of all collections were taken more than three hours before or after they were supposed to be taken. Lastly, we tested whether LB differed significantly from 0 using a t-test on LB. Table 3.3 provides a summary of our t-test on LB.

Table 3.3: T-test summary for the value of TD

|  |  |  |
| --- | --- | --- |
| **P-value** | **Mean** | **95% CI** |
| <0.01 | 2.9 | (2.87, 2.97) |

With a significant p-value, we can conclude that the actual duration of intervals between collections differed about 2.9 hours from the mandated interval. We remind the reader that this conclusion only applies to collections with a six-hour monitoring interval. It does not apply to collections with a daily AM monitoring interval.

**4. Time in Therapeutic Range**

In this section, we will address the first of two measures to evaluate patient performance: time spent within therapeutic range. There are three methods that the hospital uses to determine time spent within therapeutic range. The first is the Fraction of Values method (FOV). This method counts the number of collections within therapeutic range and divides it by the total number of collections [[2], 9]. The FOV method does not account for the variation or time between collections. It provides an equal weight to all collections regardless of the variation in aPTT or time that elapses between collections.

The second method, Stair Step (SS), account for the time but not the variation in aPTT. It assigns the time between two collections to the aPTT (Above, Inside or Below therapeutic range) of the latter collection [[2], 9]. For example, let two consecutive collections with the same episode id have aPTT values 90 (Inside) and 100 (Above) respectively, and let two hours be the time that elapses between them. Then, those two hours would count as time above therapeutic range since the latter collection has an aPTT above therapeutic range. The total time inside therapeutic range is divided by the total duration to obtain the percent of time inside therapeutic range. Recall that therapeutic range is defined as aPTT values ranging from 63 to 95.

The third method is called Linear Interpolation (LI). This method assumes that the aPTT behaves linearly. It calculates the slope between two measurements and determines the time the aPTT crosses the lower and upper bound of the therapeutic range [[2], 9]. In the previous example, we would obtain a slope (aPTT/hours) of 5. Consequently, the aPTT would cross the 95 mark an hour after the first collection. Therefore, the first hour is attributed to time in range while the second hour is attributed to time above range. Similarly, the total time inside therapeutic range is divided by the total duration to obtain the percent of time inside therapeutic range.

The linear interpolation method is considered the best approximation of time spent in therapeutic range because it accounts for both time and variation between collections. For example, episode 20’s first and second collections have aPTT values of 44.9 (Below) and 123.8 (Above) respectively. Here, FOV and SS would record no time spent in therapeutic range between measurements but LI would. Figures 4.1, 4.2, and 4.3 demonstrate this scenario. About 6.4 percent of episodes had no aPTT in therapeutic range for FOV and SS but had a nonzero value for LI.

Figure 4.1: Graph of Fraction of Value Method (Episode Id 121)

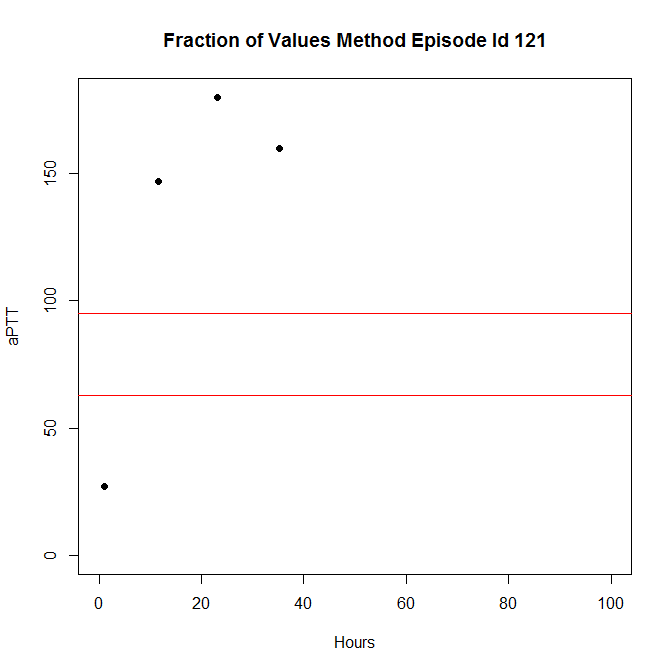


Figure 4.2: Graph of Stair Step Method (Episode Id 121)

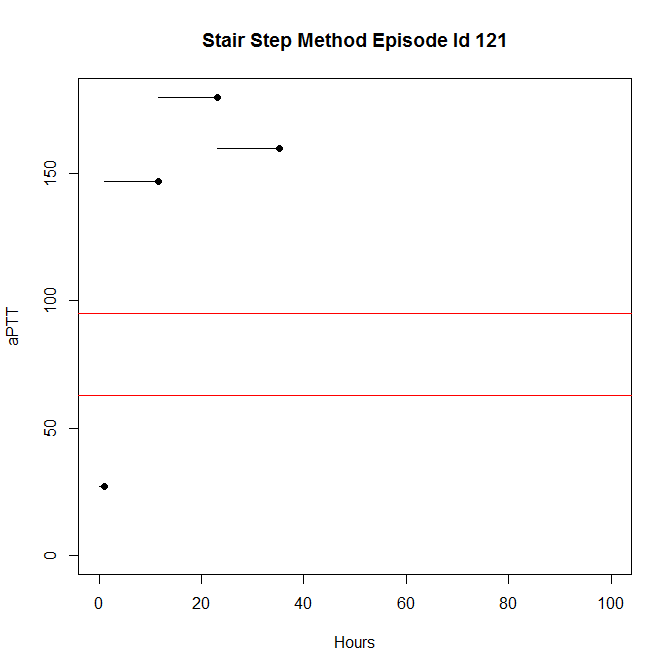
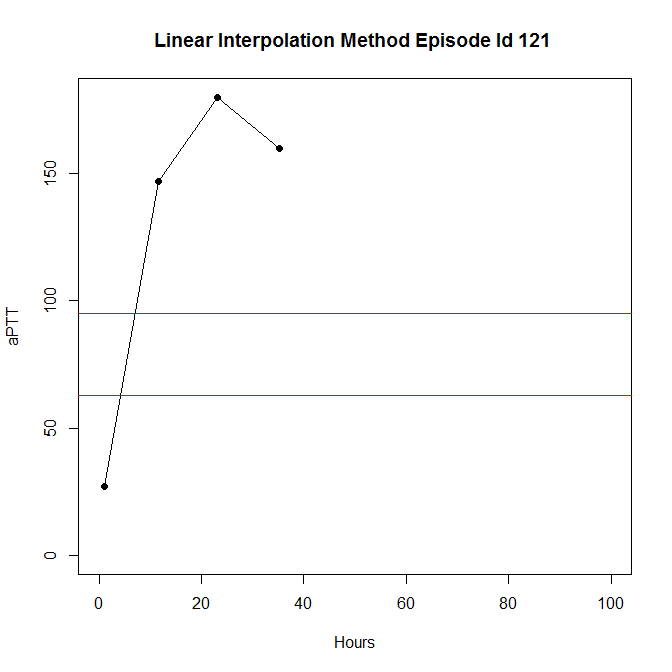


Figure 4.3: Graph of Linear Method (Episode Id 121)



It could also be that FOV and SS records time spent in therapeutic range but LI does not. This occurs when an aPTT measurement is either 63 or 95 (boundary values), but the previous and subsequent measurements were both below or both above therapeutic range. Figures 4.4, 4.5, and 4.6 demonstrate this situation.

Figure 4.4: Graph of Fraction of Value Method (Episode Id 10192)

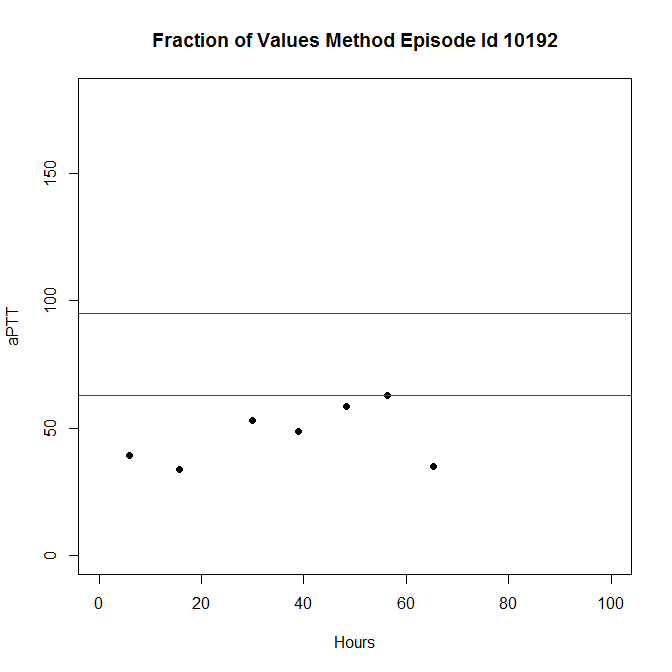


Figure 4.5: Graph of Stair Step Method (Episode Id 10192)

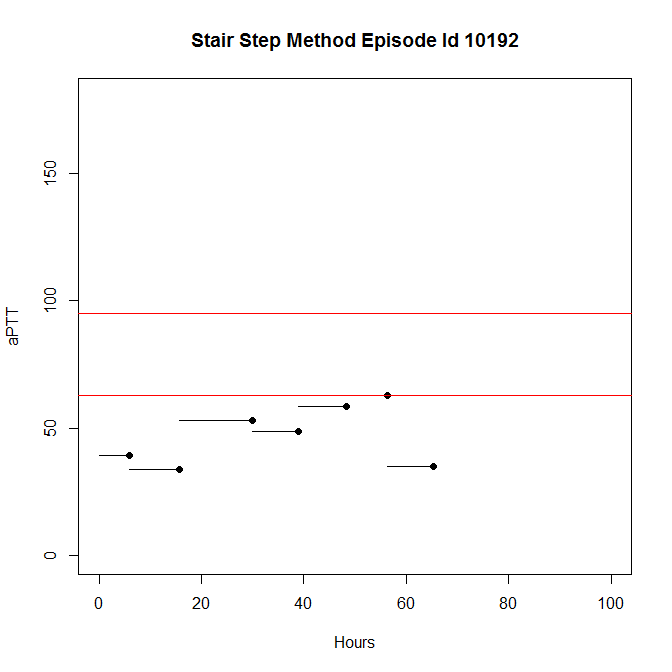
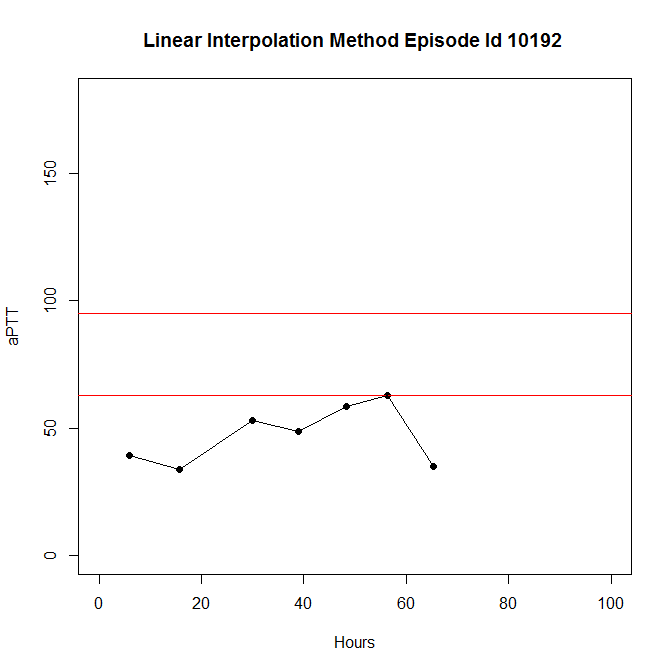


Figure 4.6: Graph of Linear Method (Episode Id 10192)



We decided to look at the distributions of time spent in therapeutic range between methods. The FOV and SS method both estimated that 84.4 and 82.8 percent of all episodes attained therapeutic range at some time or another. On the other hand, the LI method estimated that 90.7 percent of all episodes attained therapeutic range at some time or another. Table 4.7 and Figure 4.8 provide a numerical summary of the percent of time spent in therapeutic range and a boxplot of their distributions.

Table 4.7: Numerical summary of percent of time spent in therapeutic range for each method.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Method** | **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| **FOV** | 0 | 20 | 38.9 | 37.5 | 53.9 | 100 |
| **SS** | 0 | 17.8 | 36.5 | 36.9 | 54.2 | 100 |
| **LI** | 0 | 21.5 | 42.2 | 41.8 | 60.9 | 100 |

Figure 4.8: Boxplots of the time spent in therapeutic range for each method.

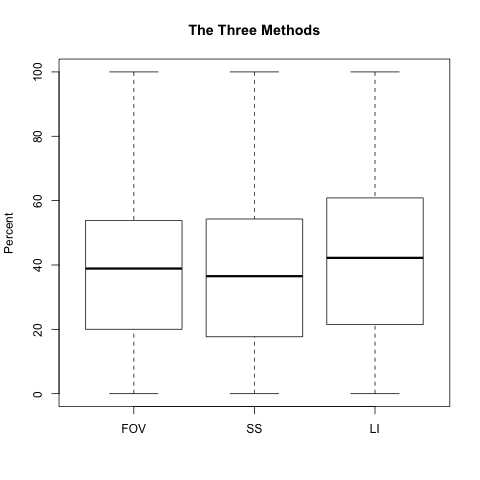


Table 4.7 shows that all methods estimate the mean percent of time spent in therapeutic range to be around 37 to 42 percent. We performed a t-test to estimate the mean percent of time inside therapeutic range. We used the percentages based on the LI method because we considered it the best estimate among all methods. Table 4.9 provides the results of our t-test.

Table 4.9: T-test for percent of time spent in therapeutic range (based on LI)

|  |  |  |
| --- | --- | --- |
| **P-value** | **Mean** | **95% CI** |
| <0.01 | 41.8 | (41.1,42.6) |

Therefore, on average, episodes spend about 41.8 percent of their time in therapeutic range. However, the median tells us that 50 percent of all episodes spend less than 42.2 percent of their time below therapeutic range.

**Comparison of Methods**

We decided to analyze whether FOV, SS and LI would yield a significantly different percentage of time in therapeutic range. Table 4.7 showed that all methods have similar distributions. This was counterintuitive to what we expected. We expected them to differ due to difference in methodology. However, we realized that a comparison of the distributions did not provide an accurate comparison of the methods; an analysis of the pairwise difference (by episode) would be more suitable. Table 4.10 provides a numerical summary for the pairwise difference between methods. Figure 4.11 provides a boxplot for the pairwise difference between methods.

Table 4.10: Numerical summary for the pairwise difference between methods

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Pairs** | **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| **FOV-SS** | -64.4 | -4.2 | 0 | -0.5 | 5.6 | 66.7 |
| **FOV-LI** | -78.6 | -12.0 | -3.5 | -4.4 | 1.4 | 61.8 |
| **SS-LI** | -84.8 | -13.0 | -3.85 | -4.9 | 1.3 | 97.2 |

Figure 4.11: Boxplot for the pairwise difference between methods

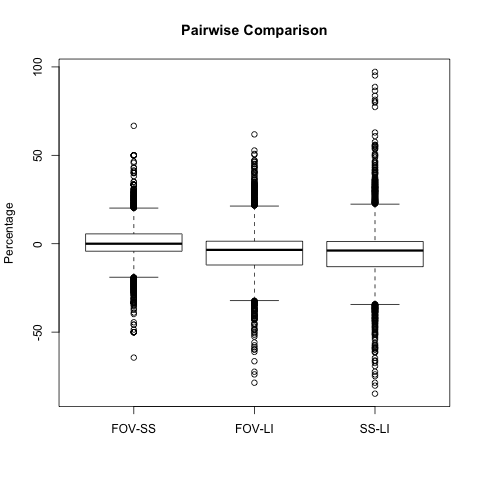


Table 4.10 suggests that these methods do differ from each other but not by much since the means are all relatively close to 0. A mean of -0.9 suggest that SS produces slightly higher percentages of time spent in therapeutic range compared to FOV. The other two means suggest that LI produces higher percentages of time spent in therapeutic range compared to the other two methods.

The maximum for the difference between SS and LI (97.2) was shocking. This was traced back to episode 14658. Episode 14658 only had two collections. The first collection was above therapeutic range (132.8) and the second was in therapeutic range (93.9). Therefore, the FOV, SS and LI returned a percent of time of 50, 100 and 2.8 respectively. This example illustrates how different each method is from each other.

Figure 4.11 shows that the difference between FOV and LI varies less than the difference between SS and LI. The standard deviations for FOV-SS, FOV-LI and SS-LI are 10.5, 12.9, and 14.9 respectively. This is also counterintuitive. We expected a larger variation between FOV and LI since FOV is considered the least accurate method.

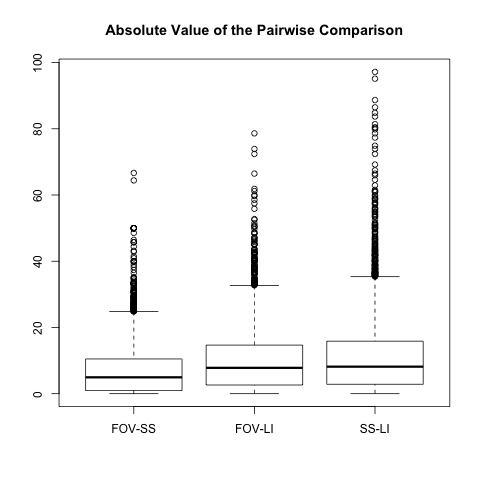
Figure 4.11 also shows a balance of positive and negative differences. This made us realize that the mean difference between methods were misleading. The means were relatively close to zero because the positive and negative values were averaging each other out; For example, the minimum and maximum difference for the FOV-LI methods are -78.6 and 61.8, respectively. When calculating the average, these values negated each other to produce a mean that was closer to zero. However, both values indicate a large difference in methodology.

To account for these negations, we considered the absolute value of the differences among methods. Table 4.12 provides a numerical summary of the absolute value of the differences in percent between methods. Figure 4.13 provides a boxplot for the absolute value of the difference in percent between methods.

Table 4.12: Numerical summary for the absolute value of pairwise differences

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Pairs** | **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| **FOV-SS** | 0.0 | 0.9 | 4.9 | 7.1 | 10.5 | 66.7 |
| **FOV-LI** | 0.0 | 2.6 | 7.8 | 9.9 | 14.7 | 78.6 |
| **SS-LI** | 0.0 | 2.8 | 8.2 | 11.1 | 15.8 | 97.2 |

Figure 4.13: Boxplots for the absolute value of the differences between methods.



The FOV and SS methods have an average absolute value difference of 7.1. Therefore, on average, the FOV and SS methods differ by 7.1 percent. Similarly, on average, the FOV and LI methods differ by 9.9 percent, and the SS and LI methods differ by 11.1 percent. This implies that there is difference of percent in therapeutic range depending on the method used.

To test whether the absolute value of the difference differed significantly from zero, we conducted a t-test on the absolute value of the difference in percentages. Table 4.14 provides the results for our t-test.

Table 4.14: Results from the t-test for each difference.

|  |  |  |  |
| --- | --- | --- | --- |
| **Pairs** | **P-value** | **Mean** | **95% C.I.** |
| **FOV-SS** | <0.01 | 7.1 | (6.9, 7.3) |
| **FOV-LI** | <0.01 | 9.9 | (9.6, 10.2) |
| **SS-LI** | <0.01 | 11.1 | (10.7, 11.3) |

The t-test was significant for all three pairs. On average, FOV and SS differ between 6.9 and 7.3 percent while FOV and SS both differed from LI by about 10 to 11 percent. This is a significant result considering that the average time in range for the LI method was about 42 percent. If the LI method did return a difference of 42 percent, the FOV and SS method could range anywhere from 30 to 50 percent, which might make a big difference as far as the patient’s performance is concerned.

**5. Time to initial aPTT**

In Section 4, we discussed the time spent inside therapeutic range as a measure of patient performance. Another measure of patient performance is the time patients take to enter therapeutic range. This is known as the time to initial aPTT. One can think about time to initial aPTT as a measure of the hospital’s ability to learn a patient’s aPTT behavior. If a patient took a long time to enter therapeutic range, it suggests that hospital staff have not understood how to control the patient’s aPTT.

Dr. Patrick details two methods to calculate time to initial aPTT. The first method defines time to initial aPTT as the time till the first recorded aPTT inside therapeutic range [[2], 10]. This is defined as the time difference between the order start and collection time of the first recorded aPTT inside therapeutic range. This was easily obtained using the “TimetoCollection” column.

The second method is based on linear interpolation. It assumes that all patients begin with an aPTT value of 29 and assumes that the aPTT behaves linearly between collections. Time to initial aPTT is defined as the time it takes a patient’s aPTT to cross the lower boundary of therapeutic range (65) [[2], 10]. Due to time constraints, we did not analyze time to initial aPTT based on the second method.

As mentioned, 84.5, 82.8 and 90.7 percent of all patients spent time in therapeutic range based on FOV, SS and LI respectively. Table 5.1 summarizes the time to initial aPTT for episode that reached therapeutic range for the first method of calculating time to initial aPTT.

Table 5.1: Numerical summary for time (hours) to initial aPTT for episodes that reached therapeutic range.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| 0.00 | 9.5 | 19.2 | 26.4 | 34.6 | 351.3 |

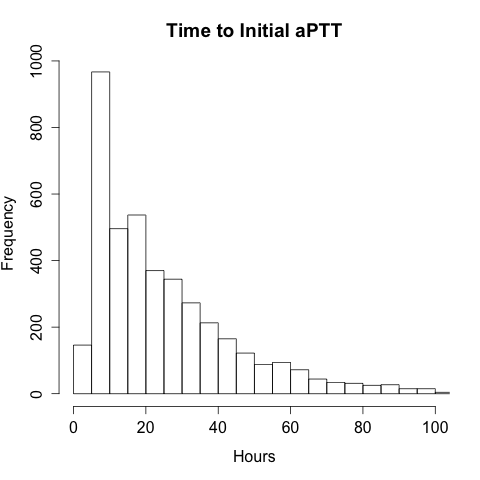
The minimum of 0 occurs when the first measurement is within the therapeutic range and the order start time and the first collection time are the same. These collections (67 collections) were addressed in Section 3 while trying to determine whether we should include or exclude them. We re-ran the numerical summary without the episodes from these 67 collections and obtained the following results. Suffice, to say it did not change much.

Table 5.2: Numerical summary for time (hours) to initial aPTT (Re-ran)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Min.** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max.** |
| 0.02 | 9.5 | 19.2 | 26.3 | 34.5 | 351.3 |

Out of the episodes that achieved initial aPTT, 1.5 percent took over 100 hours to achieve initial aPTT and 58.1 percent took less than 24 hours to achieve initial aPTT. The histogram in Figure 5.2 visually represents the number hours for the episode Ids to reach an initial aPTT. Since only 1.5 percent of episode Ids took over 100 hours, we excluded them to obtain a better visualization of our histogram.

Figure 5.2: Histogram of Time to Initial aPTT



A t-test revealed a mean of 26.3 with a 95 percent confidence interval between 25.6 and 27.0 with a significant p value. Thus, on average, episodes take about 26.3 hours to reach initial aPTT.

Dr. Patrick wanted to determine if there was a relationship between time to initial aPTT and percent spent in therapeutic range. Figure 5.3 and Figure 5.4 provide scatterplots of the time to initial aPTT and the percent in therapeutic range based on the LI method. Figure 5.4 only plots points that took less than 100 hours to achieve initial aPTT.

Figure 5.3: Scatterplot of time to initial aPTT vs. Percent in therapeutic range.

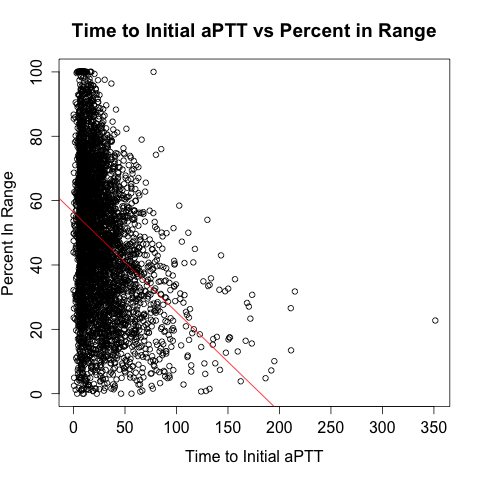
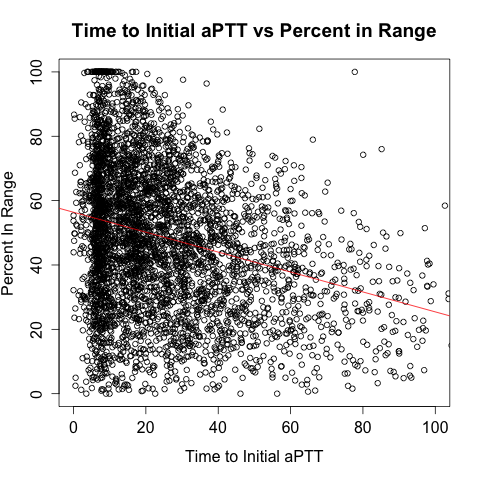


Figure 5.4: Scatterplot of time to initial aPTT vs. Percent in therapeutic range.



Based on the scatterplots, there does not seem to be a strong linear relationship for time to initial aPTT and percent in therapeutic range. We fitted a linear model and obtained a model () with an adjusted R-squared of 0.11 and a significant p-value. Thus, time to initial aPTT only accounts for about 11 percent of the variation for percent spent in therapeutic range. Therefore the time to initial aPTT does not correlate with the percent of time spent inside therapeutic range.

In his paper, Dr. Patrick defines another measure of hospital performance called time to stable aPTT. It is defined as the time a patient takes to achieve consecutive aPTTs inside therapeutic range. It can be estimated using the same two methods used to estimate time to initial aPTT. However, we did not perform any analysis on time to stable aPTT due to time constraints.

**6. Conclusion**

In Section 2, we provided an overview of the dataset given by Dr. Patrick. We discussed the structure of the dataset, the modifications that were made to the dataset and the additional dataset that was created using the original dataset.

We also discussed the presence of dataset anomalies in Section 2. We found collections with identical episode Id and collection times as the previous collection but with different numerical results. We also found collections where the order start time was the same as the first collection time of the episode. In total, 917 collections from 50 episodes were excluded from our study. The second type of anomaly was only excluded from our analysis in Section 5 because they did not affect our results in Section 3 and 4.

In Section 3, we discussed the hospital’s performance to follow hospital procedure. We only performed an analysis on collections with a six-hour monitoring interval due to time constraints. A t-test revealed that on average, the actual collection time (for collections with a six-hour monitoring interval) was late by about 2.9 hours.

We mentioned that we were unable to perform a similar analysis on collections with a daily AM monitoring interval due to time constraints. This is an area worth looking into. It would be interesting to see whether the hospital performs differently for daily AM monitoring interval compared to six-hour monitoring intervals.

A survival analysis is another suitable analysis that we did not carry out. Theoretically, we would like to quantify the likelihood of the subsequent collection being taken between certain periods of time. We would treat the collections as event and the time between collections as the independent variable. This would require some restructuring of the dataset, but it is definitely feasible. A survival analysis could also be carried out on the time to initial aPTT.

In Section 4, we introduced time in therapeutic range as a measure of hospital performance and three different methodologies (FOV, SS and LI) to measure it. We discussed why LI was a more superior method compared to the other two methods and used calculations based on LI to show that on average, episodes spend about 41.8 percent of their time in therapeutic range. However, the median showed that less than 50 percent of all episodes spend less than 42.2 percent of their time in therapeutic range.

We also performed a comparison between methodologies in Section 4. We found that on average, the percent of time spent in therapeutic range differed by 7.1 percent between FOV and SS, 9.9 percent between FOV and LI and 11.1 percent between SS and LI.

In Section 5, we introduced another measure of hospital performance – time to initial aPTT – and two different methodologies to calculate it. However, due to time constraints, we only managed to perform an analysis on one of the methods. We found that on average, episodes took about 26.3 hours to reach initial aPTT.

We also analyzed the relationship between time to initial aPTT and time spent inside therapeutic range. Our model showed that time to initial aPTT only accounted for 11 percent of the variation of time spent in therapeutic aPTT. Therefore, time to initial aPTT was not a good predictor of time spent in therapeutic aPTT.

We have yet to write code to calculate time to initial aPTT based on the linear interpolation method. Once written, this would allow us to re-analyze the relationship between time to initial aPTT and time spent in therapeutic aPTT. We could then compare the results of this model with the earlier model to examine how much they differ from each other.

We briefly introduced the “Bolus” column in Section 2. However, we did not use it in our analysis. An initial bolus is given at the beginning of treatment if the aPTT is exceptionally low. Dr. Patrick wanted to examine the relationship between the presence of an initial bolus and the time to initial aPTT. He wanted to know whether the presence of an initial bolus decreased the time to initial aPTT. This can be examined using a one-way ANOVA model with the initial bolus as the independent variable and the time to initial aPTT as the dependent variable.

We realize there is a limited amount of graphical analysis in our paper. It was challenging to produce simple graphics that would capture the variation in aPTT and TD across multiple episodes. Toward the end of this paper, we wanted to recreate the change in time spent inside therapeutic range that Dr. Patrick created. However, we did not have enough time to do so.

Another method of visualizing our data is to plot the TD column against the time of the day. By extracting the hour (or alternate desired unit) of the day from the “Collection\_Time” column, one could theoretically plot the mean TD against the specific time interval of the day. This would allow us to graphically see how hospital staff performance (being on time to perform collections) varied throughout the day. It would help analyze whether hospital staff perform better during specific time periods of the day.

This concludes the work we have done thus far. We want to thank Dr. Patrick for letting us collaborate with him on this project. This experience was challenging, thought provoking and fulfilling. We hope this paper provided the reader with a better understanding of the hospital’s performance for anticoagulation with IVUFH. We also hope that it inspires the hospital to find creative solutions to better improve its performance in order to provide excellent healthcare to the patients they are serving.

**7. References**

[1] Grolemund, Garrett., and Hadley Wickham. “Dates and Time Made Esasy with {lubridate}.” *Journal of Statistical Software 40.3* (2011): 1-25. <http://www.jstatsoft.org/v40/i03/>

[2] Patrick et al. “Measuring Hospital Performance for Anticoagulation with Intravenous Unfractionated Heparin : A Comparison of Methods.” : 1-10 (Unpublished)

[3] R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing, 2015. <http://www.R-project.org/>.

[4]Wickham, Hadley. “The Split-Apply-Combine Strategy for Data Analysis”.

*Journal of Statistical Software 40.1* (2011) : 1-29 <http://www.jstatsoft.org/v40/i01/>